www.bsc.es

Madrid, 26 abril 2017

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

ESiWACE

Excellence in Simulation of Weather and Climate in Europe

Kim Serradell – Computational Earth Sciences

Indice

- Introducción
- Objetivos
- Modelos de ciencias de la tierra
- Actividades realizadas en el BSC

¿Qué es ESiWACE?

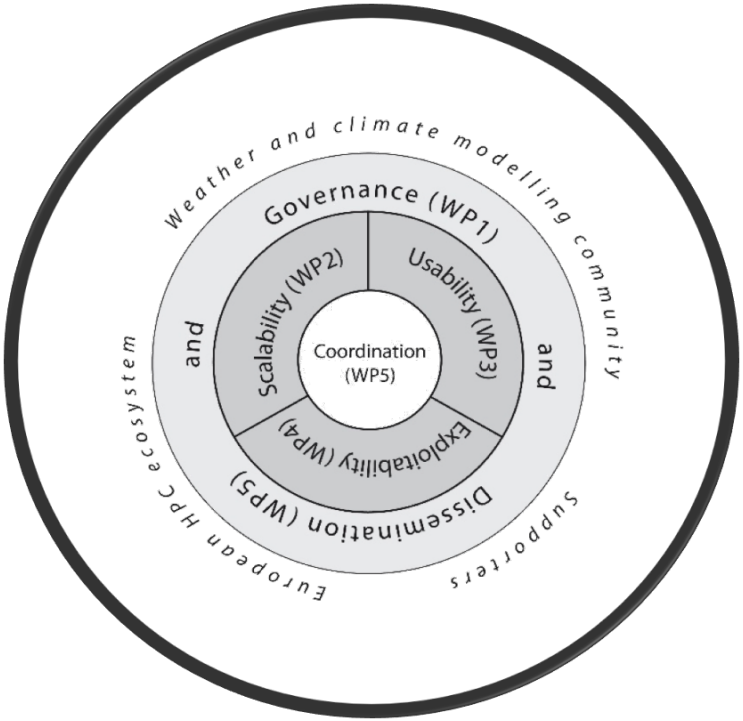
- Call: EINFRA-5-2015: Centres of excellence for computing applications
- Kick-off: 1 de diciembre de 2015 (dos revisiones ya realizadas)
- Duración: 48 meses
- Participantes (en orden de PMs): **DKRZ (líder)**, ECMWF, CNRS-IPSL, MPG, CERFACS, BSC, STFC, MET Office, U. Reading, SMHI, ICHEC, CMCC, DWD, SEAGATE, Bull and ALLINEA.

¿Porqué ESiWACE?

- La comunidad de meteorología y clima tiene una **"casi infinita"** necesidad de recursos de computación
  - Más resolución horizontal
  - Mayor número de miembros en los "ensembles"
  - Representar más procesos en los modelos

- Algunos de nuestros problemas con HPC
  - Aparecen nuevas tecnologías pero ello no implica que nuestros modelos corran mejor
  - Dificultad creciente en muchos aspectos (arquitecturas, modelos…)
  - Códigos con años de funcionamiento en una comunidad tradicionalmente reacia a los cambios
  - Grandes volúmenes de datos
  - Tenemos poca capacidad de influencia en el mercado comercial de HPC

Organización de ESiWACE



WP1 Governance and engagement

- Engagement and governance
- Enhancing community capacity in HPC
- Strategic interaction with HPC ecosystem and HPC industry
- Sustainability and business planning

WP2 Scalability

- Support, training and integration of state of-the-art community models and tools
- Performance analysis and inter-comparison
- Efficiency enhancement of models and tools
- Preparing for exascale

WP3 Usability

- ESM end-to-end workflows recommendations
- ESM system software stack recommendations
- ESM scheduling
- Co-Design for Usability

WP4 Exploitability

- The business of storing and exploiting high volume climate data
- New storage layout for Earth system data
- New methods of exploiting tape
- Semantic mapping between netCDF and GRIB

WP5 Management and Dissemination

Dos grandes objetivos

- Objetivo científico
  - Modelo global *"cloud resolving"*
    - El objetivo principal es correr modelos globales con resoluciones espaciales que permitan simular *"convective clouds"* y *"small-scale ocean eddies"*. Estas simulaciones producirán mejores predicciones y representaran mejor procesos a escala regional.

- Objetivo técnico
  - Modelo global a 1km de resolución con un rendimiento eficiente
    - Necesitamos resoluciones espaciales de 1 km (o mayores). Idealmente, se deben producir entre 100 y 1000 días de simulación al día.

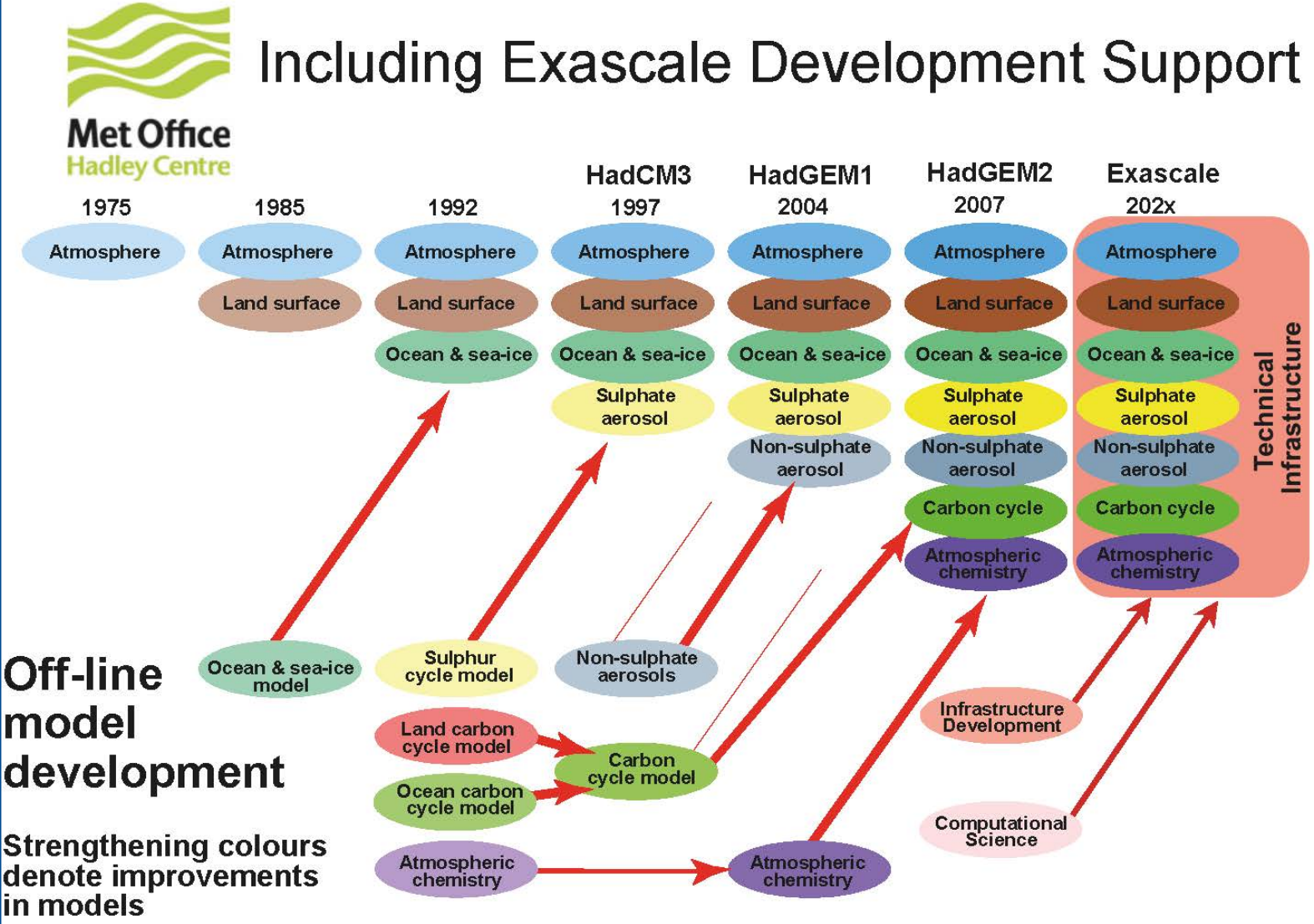
www.bsc.es

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

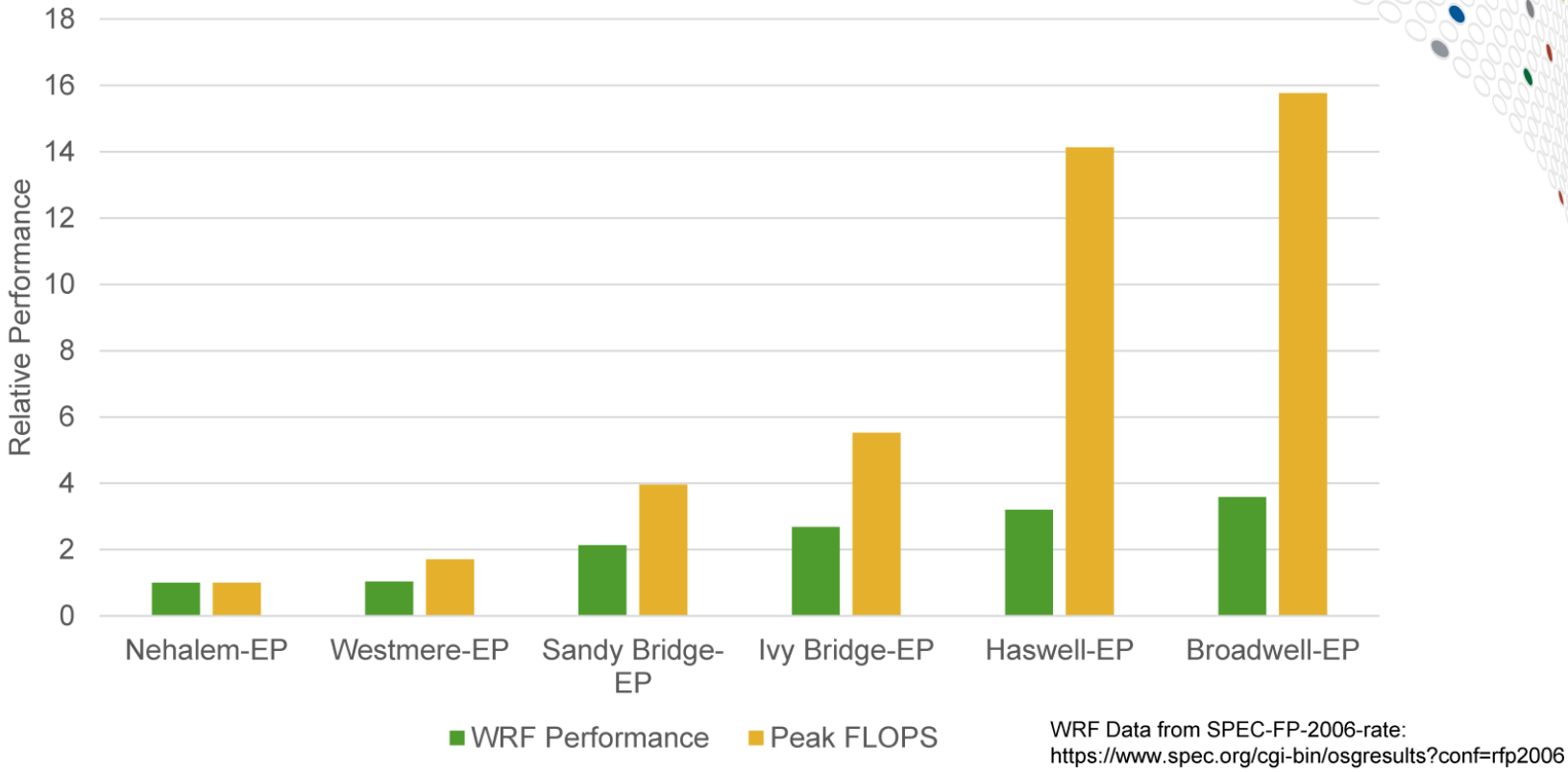
EXCELENCIA SEVERO OCHOA

Modelos de ciencias de la tierra

Sistemas complejos…



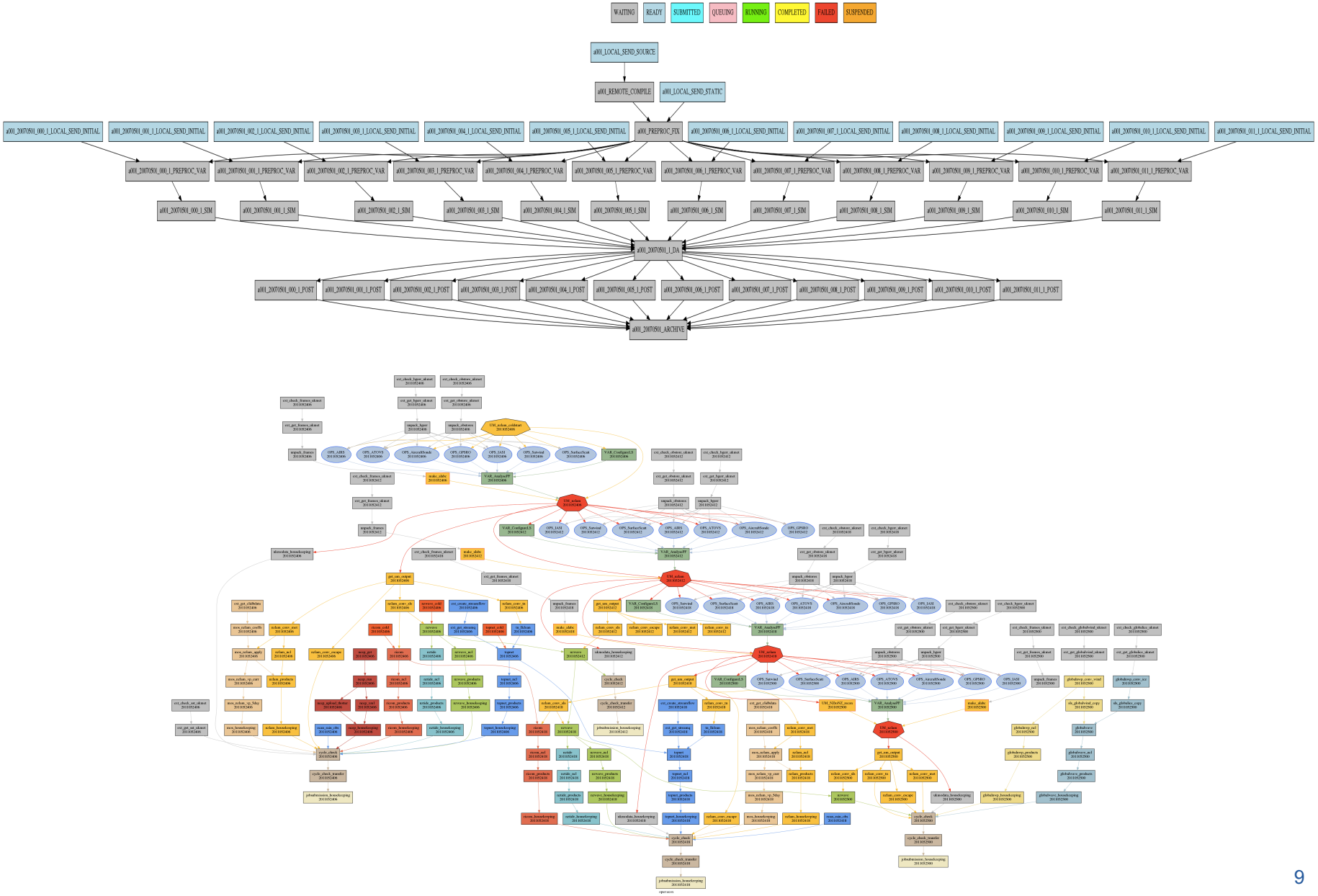
FLOPs aren't the bottleneck!



WRF Data from SPEC-FP-2006-rate:
https://www.spec.org/cgi-bin/osgresults?conf=rfp2006

Copyright 2016 Cray Inc.

… y difíciles de manejar



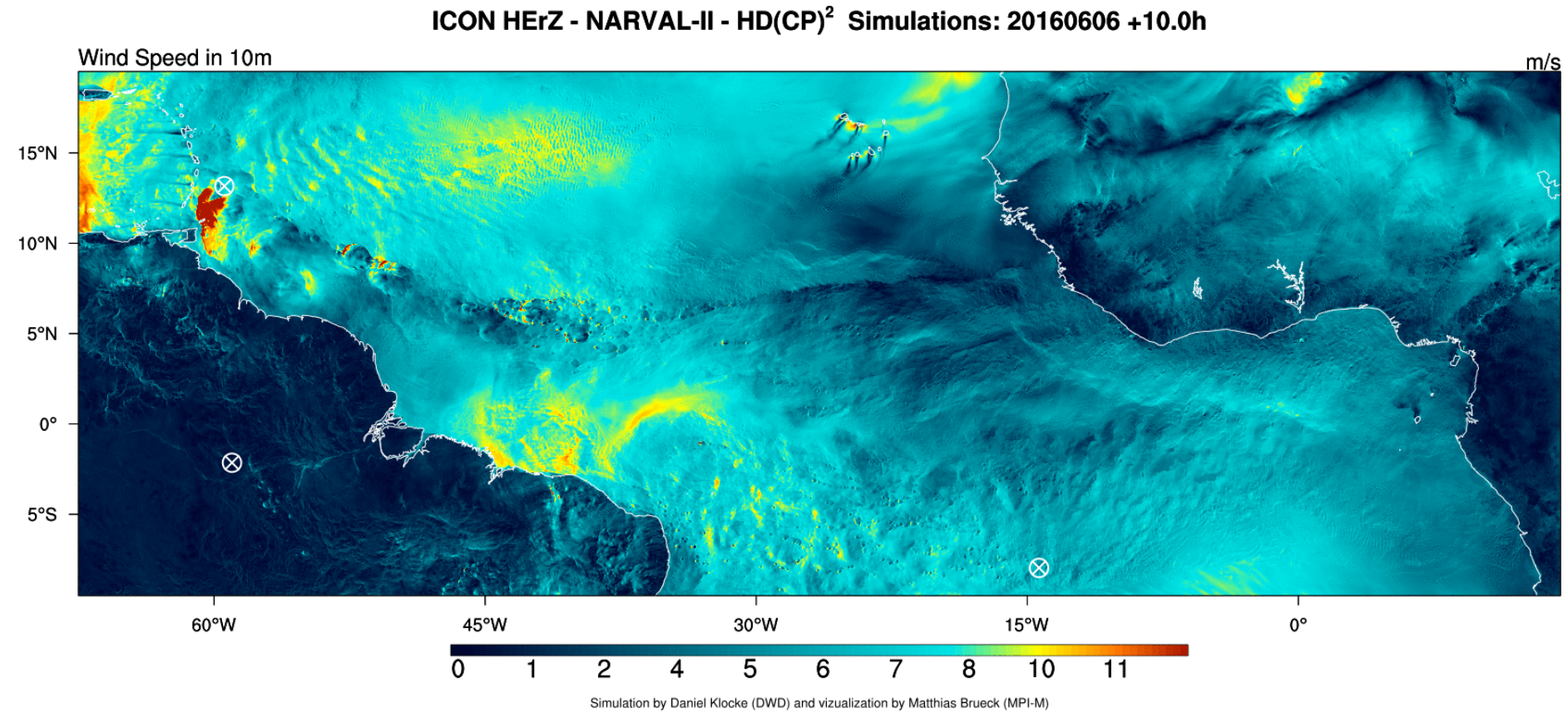
“Demonstrators”

- ESiWACE quiere demostrar que es factible correr:
  - 1km global atmosfera, IFS, 2017
  - 1km global atmosfera, ICON, 2017
  - 1km global océano, NEMO, 2019
  - al menos 10km global, acoplado ESM, EC-EARTH, 2019
  - 1km global, acoplado ESM, MPI-ESM2, 2019

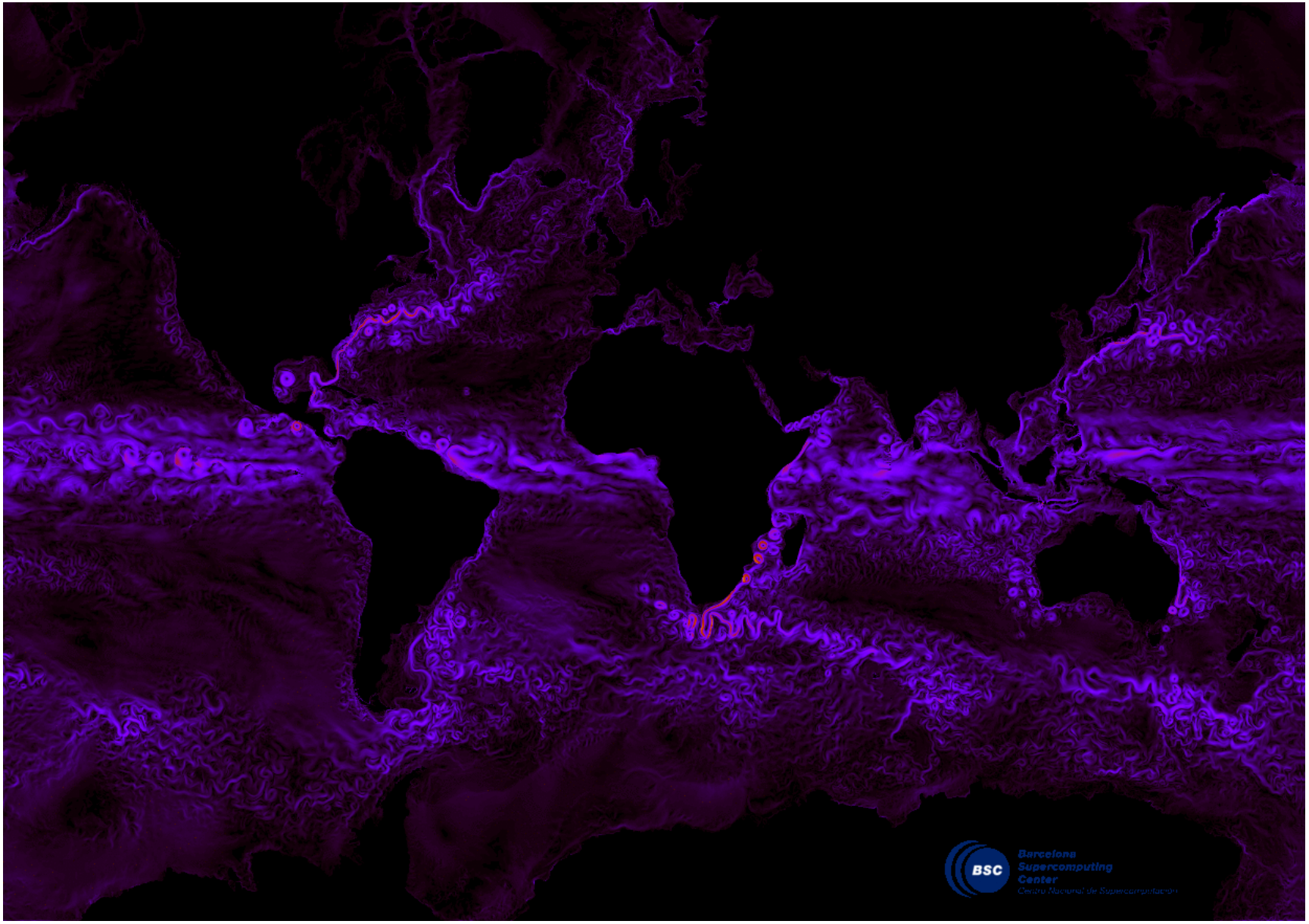
- El reto y la diversidad del problema justifican un proyecto coordinado para su investigación.

- Desarrollar la próxima generación de códigos meteorológicos y de clima de cara al Exascale, es un proyecto a largo plazo que va más allá de ESiWACE.

ICON 1km


ICON HErZ - NARVAL-II - HD(CP)² Simulations: 20160606 +10.0h
Wind Speed in 10m
m/s
15°N
10°N
5°N
0°
5°S
60°W
45°W
30°W
15°W
0°
0 1 2 4 5 6 7 8 10 11
Simulation by Daniel Klocke (DWD) and vizualization by Matthias Brueck (MPI-M)

Modelo climático acoplado a 10km



www.bsc.es

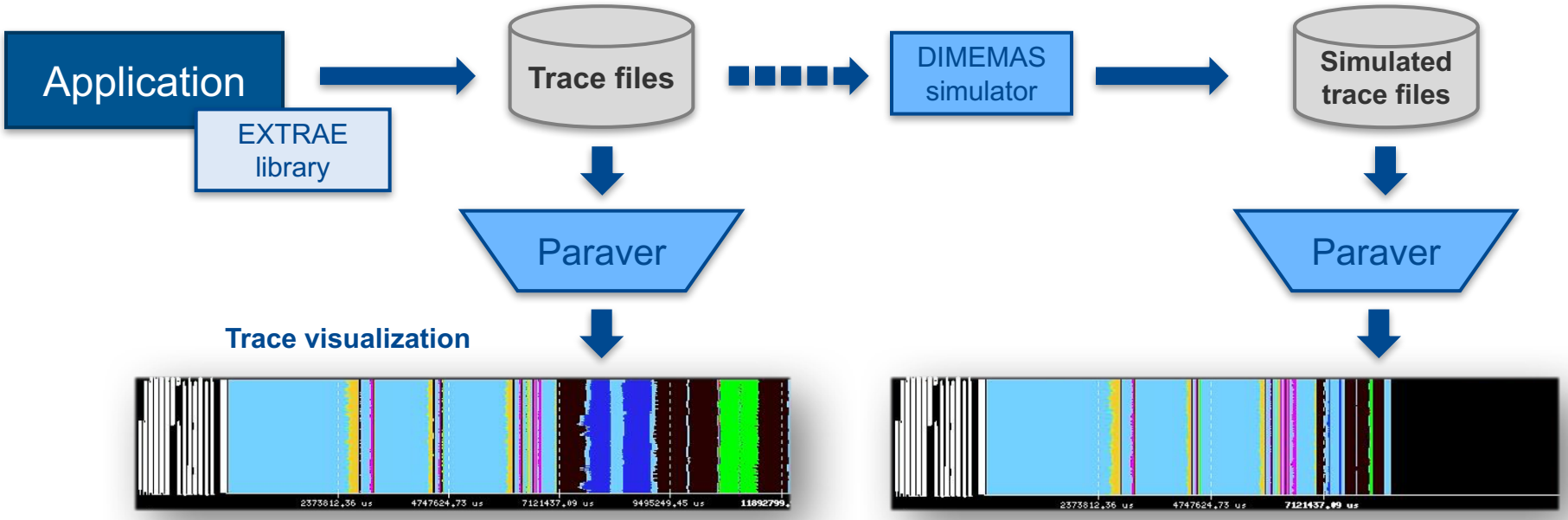
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

Nuestra aportación en ESIWACE

Escalabilidad

- Since 1991
- Based on traces
- Open Source: http://www.bsc.es/paraver
- **Extrae**: Package that generates Paraver trace-files for a post-mortem analysis
- **Paraver**: Trace visualization and analysis browser
  - Includes trace manipulation: Filter, cut traces
- **Dimemas**: Message passing simulator



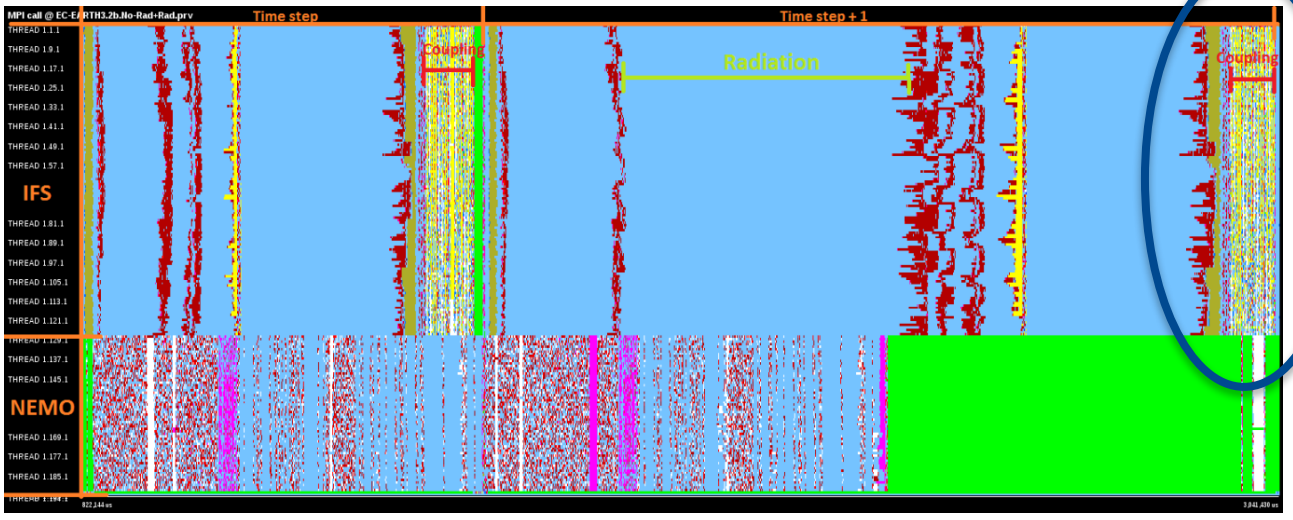
DIMEMAS generated trace. Target = ideal machine

Escalabilidad (II)

- Análisis del rendimiento del modelo EC-Earth

IFS: 128 cores

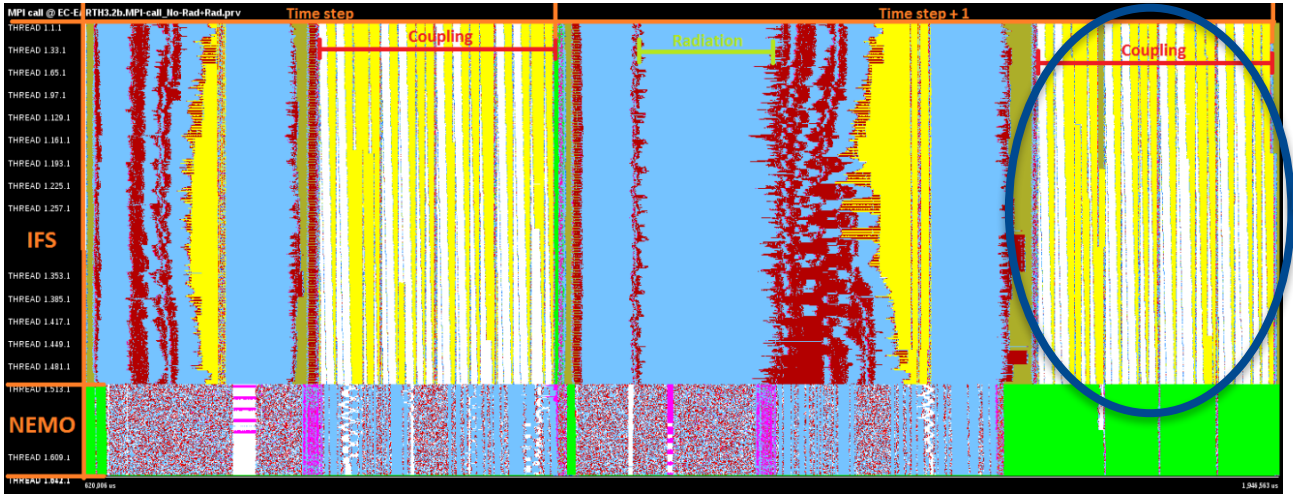
NEMO: 128 cores



tiempo

IFS: 512 cores

NEMO: 128 cores



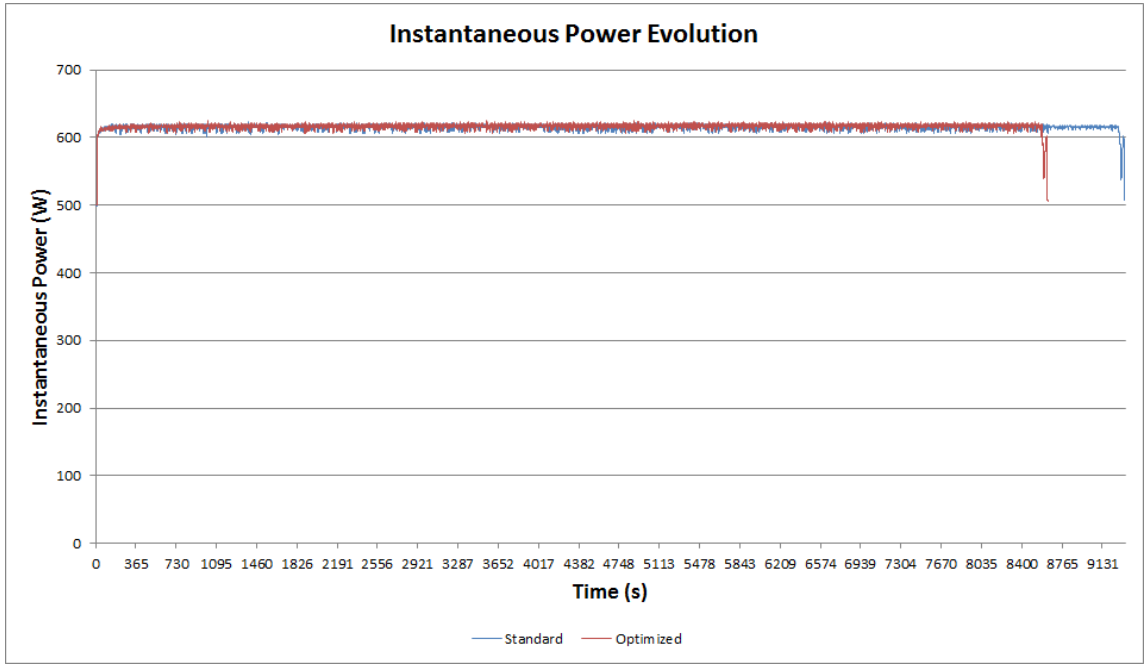
Incremento en el tiempo del couple

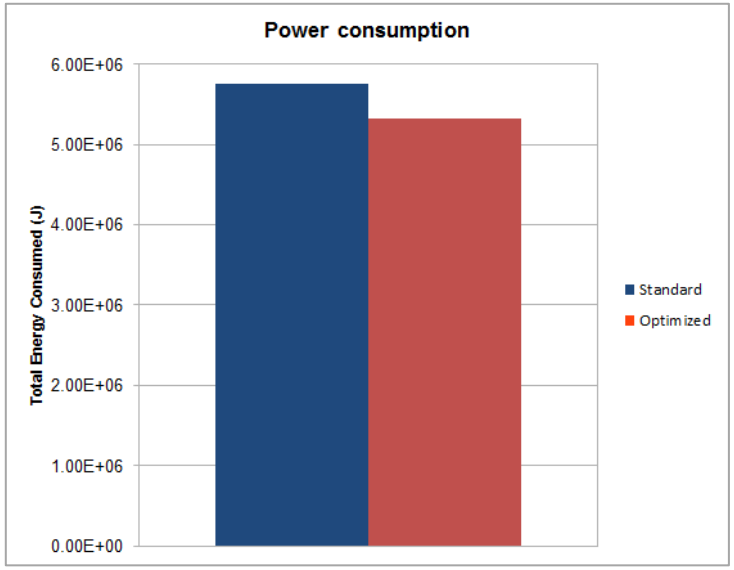
Escalabilidad (III)

- Reducción del tiempo de acoplamiento en un 80%
- Usando la configuración estándar, el tiempo de ejecución **se reduce en un 23%**
- Estas mejoras ya están en distribuidas y usadas por todos los miembros de EC-Earth

Portabilidad

- Portabilidad del modelo a otras arquitecturas
  - En colaboración con el proyecto Montblanc, se ha ejecutado el modelo en procesadores ARM y se ha realizado un perfil del consumo energético del modelo EC-Earth

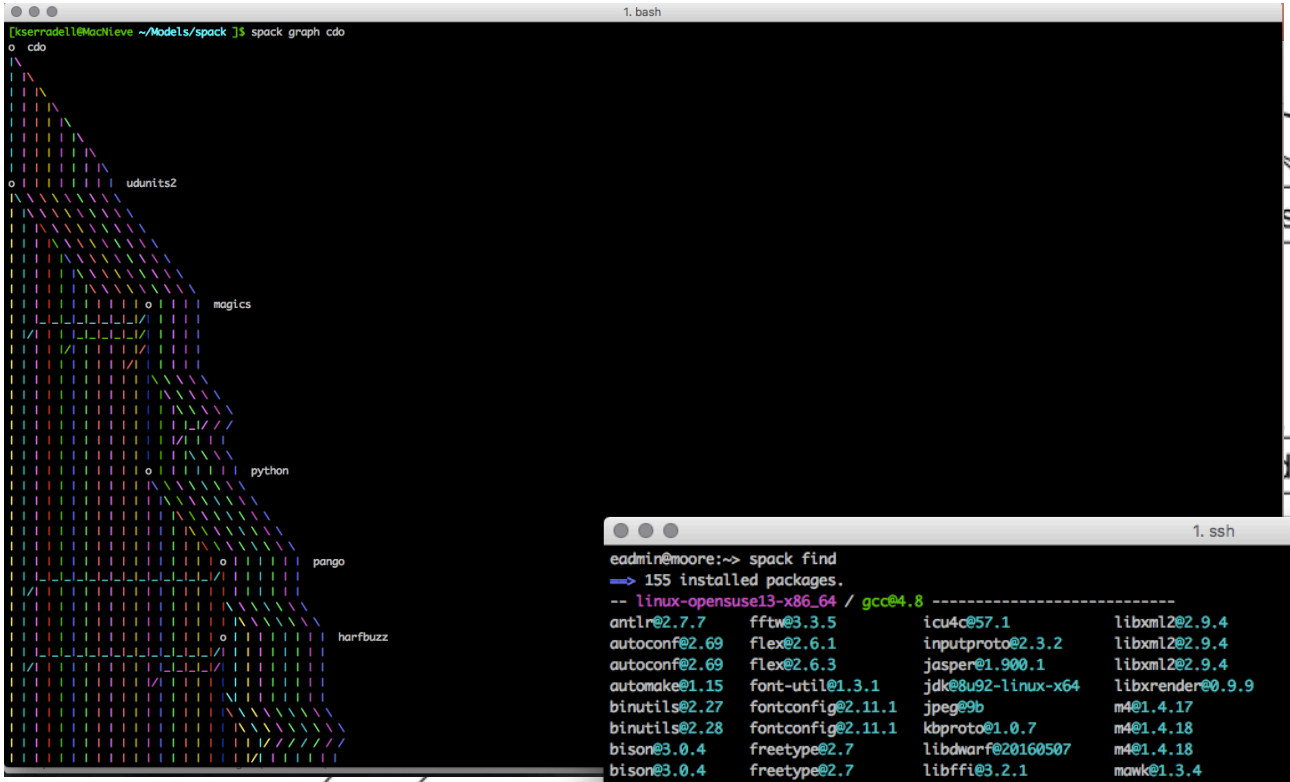


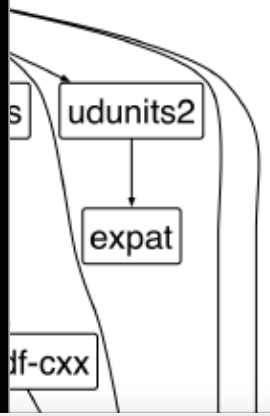


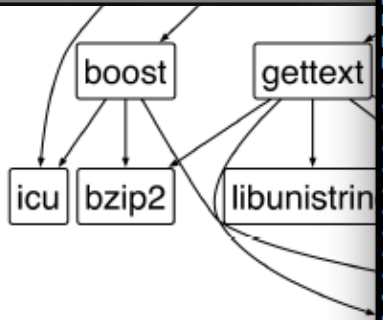
Usabilidad

- Nuestros modelos tienen muchas dependencias
  - Compiladores
  - Librerías
  - Herramientas de postproceso
  - …
- Desplegar un modelo en una nueva máquina no es fácil y toma tiempo
- Solución: automatizar el proceso
- SPACK: gestor de software orientado a la supercomputación
  - https://github.com/LLNL/spack
  - Desplegado en varias máquinas de la RES

Usabilidad (II)







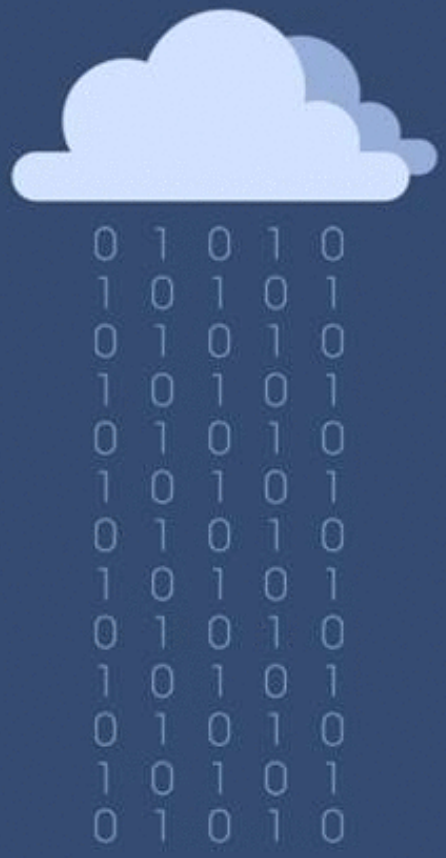
1. ssh

```
eadmin@moore:~> spack find
==> 155 installed packages.
-- linux-opensuse13-x86_64 / gcc@4.8 ----------------------------
antlr@2.7.7     fftw@3.3.5          icu4c@57.1          libxml2@2.9.4       pango@1.40.3        qhull@2015.2
autoconf@2.69   flex@2.6.1          inputproto@2.3.2    libxml2@2.9.4       papi@5.5.1          r@3.3.2
autoconf@2.69   flex@2.6.3          jasper@1.900.1      libxml2@2.9.4       parallel-netcdf@1.8.0 r-ncdf4@1.15
automake@1.15   font-util@1.3.1     jdk@8u92-linux-x64  libxrender@0.9.9    pcre@8.39           readline@6.3
binutils@2.27   fontconfig@2.11.1   jpeg@9b             m4@1.4.17           perl@5.24.1         renderproto@0.11.1
binutils@2.28   fontconfig@2.11.1   kbproto@1.0.7       m4@1.4.18           pixman@0.34.0       sqlite@3.8.5
bison@3.0.4     freetype@2.7        libdwarf@20160507   m4@1.4.18           pixman@0.34.0       szip@2.1
bison@3.0.4     freetype@2.7        libffi@3.2.1        mawk@1.3.4          pkg-config@0.29.1   tar@1.29
boost@1.63.0    gettext@0.19.8.1    libgcrypt@1.6.2     munge@0.5.11        proj@4.9.2          tcl@8.6.5
bzip2@1.0.6     gettext@0.19.8.1    libgpg-error@1.21   nasm@2.11.06        py-cdo@1.3.2        tcl@8.6.5
bzip2@1.0.6     gettext@0.19.8.1    libjpeg-turbo@1.5.0 nco@4.6.2           py-cycler@0.10.0    tk@8.6.5
cairo@1.14.0    glib@2.49.4         libpciaccess@0.13.4 ncurses@6.0         py-cython@0.23.5    tk@8.6.5
cairo@1.14.0    glib@2.49.4         libpng@1.6.26       ncurses@6.0         py-dateutil@2.5.2   udunits2@2.2.20
cairo@1.14.8    grib-api@1.17.0     libpng@1.6.27       netcdf@4.4.1        py-mock@1.3.0       util-macros@1.19.0
cairo@1.14.8    gsl@2.3             libpthread-stubs@0.3 netcdf@4.4.1       py-netcdf@1.2.3.1   uuid@1.6.2
cdo@1.7.2       harfbuzz@0.9.37     libsigsegv@2.10     netcdf@4.4.1        py-nose@1.3.7       xcb-proto@1.12
cmake@3.7.1     harfbuzz@0.9.37     libsigsegv@2.11     netcdf@4.4.1.1      py-numpy@1.11.2     xerces-c@3.1.4
cmake@3.7.2     harfbuzz@0.9.37     libtiff@4.0.6       netcdf-fortran@4.4.4 py-numpy@1.11.2    xextproto@7.3.0
cmor@3.2.0      hdf5@1.8.18         libtiff@4.0.6       netcdf-fortran@4.4.4 py-pbr@1.8.1       xproto@7.0.29
curl@7.50.3     hdf5@1.10.0-patch1  libtool@2.4.6       openblas@0.2.19     py-pillow@3.2.0     xtrans@1.3.5
curl@7.52.1     hdf5@1.10.0-patch1  libunwind@1.1       openmpi@2.0.1       py-pyparsing@2.0.3  xz@5.2.2
dyninst@9.3.0   hdf5@1.10.0-patch1  libx11@1.6.3        openmpi@2.0.1       py-pytz@2016.6.1    xz@5.2.3
elfutils@0.163  help2man@1.47.4     libxau@1.0.8        openmpi@2.0.2       py-scipy@0.18.1     zlib@1.2.8
elfutils@0.168  hwloc@1.11.4        libxcb@1.12         openssl@1.0.2j      py-setuptools@25.2.0 zlib@1.2.10
expat@2.2.0     hwloc@1.11.5        libxdmcp@1.1.2      openssl@1.0.2k      py-six@1.10.0       zlib@1.2.11
extrae@3.4.1    hwloc@1.11.6        libxext@1.3.3       openssl@1.1.0c      python@2.7.12
eadmin@moore:~>
```

“Demonstrator” de clima

- Desarrollo desde cero de la configuración de 10 kms
  - Modificación de los códigos
  - Creación de los ficheros de inicio
  - Ejecuciones en Mare Nostrum
- Se han simulado dos años en Mare Nostrum 3
  - Usando unos 2000 cores conseguimos 0.16 SYPD (años simulados por día)
- Aún queda trabajo por hacer:
  - Análisis y mejora del rendimiento (en MN4)
  - Se actualizará el modelo de atmósfera
    - IFS36r4 a oIFS43r1
  - Integración XIOS en oIFS

Preguntas



www.bsc.es

Gracias

kim.serradell@bsc.es